

APLICACIÓN DA ANÁLISE DIALECTOMÉTRICA AOS
DATOS DO *ATLAS LINGÜÍSTICO GALEGO**

Rosario Álvarez
Francisco Dubert García
Xulio Sousa Fernández

Instituto da Lingua Galega (USC)

En 1981 o profesor Hans Goebel definiu a dialectometría como unha metodoloxía que nace da alianza entre a xeografía lingüística e as técnicas de taxonomía numérica. A xeografía lingüística atende o estudo da variación lingüística no espazo e válese fundamentalmente da información ofrecida nos atlas lingüísticos, que son unha fonte de información moi rica para o coñecemento profundo dos trazos fonéticos, morfolóxicos, sintácticos e léxicos dun conxunto de lugares dun territorio. Ademais dos datos particulares sobre cada un dos puntos enquisados, os atlas permiten analizar a regularidade con que se distribúen no espazo as diferentes unidades lingüísticas, unha información que os dialectólogos consideran crucial, xa que o descubrimento desas regularidades permite recoñecer as áreas do territorio que comparten trazos lingüísticos. O procedemento empregado para descubrirlas parte da comparación dos mapas e da identificación das isoglosas que aparecen superpostas, pero topa co problema de que as coincidencias no discorrer das isoglosas son infrecuentes. En efecto, a pesar do que nos dan a entender moitos dialectólogos, os límites que marcan as fronteiras de delimitación de dúas áreas que presentan distintos trazos lingüísticos en moi poucas ocasións coinciden con exactitude, pois, como sinalaba Jakob Grimm, cada palabra, cada trazo lingüístico, ten a súa historia particular. Con todo, este feito non nos debe facer renunciar a explorar a existencia de áreas dun territorio lingüístico que

* Agradecemos a colaboración de Mónica Ares Ameijide e Marta Negro Romero na introdución e manipulación dos datos do *ALGa*.

se poidan considerar lingüísticamente uniformes, ou, para sermos máis precisos, áreas en que se dá algún grao de coincidencia na distribución dun conxunto de trazos lingüísticos.

Os procedementos da dialectoloxía tradicional resultan moi limitados para descubrir estas coincidencias. Se probamos a trazar isoglosas a feito a partir das cartas dun atlas, o resultado será unha maraña inescrutable na que as coincidencias quedan ocultas aos nosos ollos. Por iso o método de trazado de isoglosas só é posible aplicalo a un conxunto reducido de mapas e sempre facendo unha escolla previa daqueles en que se dea coincidencia na distribución de trazos. En consecuencia, de todo o material que nos ofrecen os atlas unicamente se aproveitan uns poucos mapas e desbótase unha grande cantidade de información moi valiosa.

Os estudos de variación xeolingüística recibiron nas últimas décadas unha renovación metodolóxica mercede á utilización de procedementos estatísticos para a análise dos datos dialectais, que complementan o procedemento tradicional de análise cualitativa e á vez permiten o tratamento global dun elevado número de datos. A utilización de métodos cuantitativos no estudo da variación lingüística estivo durante moito tempo restrinxido ao ámbito da sociolingüística, pois nesta disciplina das ciencias sociais nunca se cuestionou a pertinencia da estatística para a interpretación da información sobre os usos lingüísticos; en cambio, no campo da dialectoloxía os investigadores aínda teñen moitos reparos en valerse dos métodos estatísticos para comprender a variación xeolingüística.

OBXECTOS	VARIABLES		
	1320 <i>folia do piñeiro</i>	1347 <i>machada</i>	1383 <i>bolboreta</i>
C1 (O Barqueiro)	pilro	brosa	avelaíña
C2 (Chímparra)	caruma	machado	avelaía
C3 (San Claudio)	agulla	machado	mariposa
C4 (A Carreira)	pluma	machado	avelaíña
C5 (San Ramón)	agulla	machado	avelaíña

Cadro 1. Matriz de datos do *ALGa*: localidades (obxectos), preguntas (variables) e respostas (variantes)

A aplicación da estatística no estudo da xeolingüística supón tratar de xeito global o conxunto de datos dunha enquisa ou dun atlas lingüístico cunha finalidade clasificatoria. No caso dos estudos xeolingüísticos os *obxectos* analizados son os puntos de enquisa dun territorio, as *variables* son cada unha das preguntas sobre as que ten información

neses puntos e as posibles respostas para cada pregunta son as *variantes*. A análise multivariante é a técnica estatística que mellor se adapta ao tratamento de datos xeolingüísticos, xa que permite analizar de xeito global un conxunto de obxectos que están caracterizados con respecto a unha serie de variables. Este tipo de análise fundaméntase no cálculo da similaridade e da diferenza que se dá entre os obxectos estudados, e no ámbito da xeolingüística permite medir a distancia lingüística entre os puntos do territorio analizado. Asemade este método axuda a ter unha visión máis completa da variación lingüística nunha área determinada, xa que permite avaliar conxuntamente toda a información dispoñible.

O procedemento que se seguiu para o tratamento dialectométrico dos datos do *ALGa* é o deseñado por Hans Goebel para a aplicación do programa *VisualDialectometry*. Este procedemento foi aplicado na explotación de datos de varios atlas lingüísticos (*Atlas Linguistique de la France*, *Atlante Italo-Svizzero*, *Atlante Linguistico del Ladino Dolomitico e dei dialetti limitrofi*, *Atlas of England*) e parte da translación dos datos dun atlas a unha matriz de datos. Na matriz de datos (cadro 1) ordénanse as respostas do cuestionario (variantes) correspondentes a cada punto (obxecto) e pregunta (variable).

O material do *ALGa* conforma unha base de 3353 variables a partir dun cuestionario de 2711¹ preguntas realizado nunha rede de 167 localidades do territorio lingüístico. Neste traballo empregamos 321 preguntas, tanto do material xa editado como do aínda inédito (vid. anexo), que ao noso xuízo conforman un conxunto representativo da variación lingüística do galego. As preguntas escollidas son as que consideramos que mellor serven para dar conta das particularidades fonéticas, morfolóxicas e léxicas do territorio analizado; para obtermos unha análise precisa e completa da variación lingüística rexeitamos as cuestións que presentaron moitas respostas baleiras, pois poderían distorsionar os resultados².

¹ Na actualidade o número de preguntas que figura na base de datos do *ALGa* é de 3353 debido a que moitas das preguntas do cuestionario orixinal tiveron que ser repartidas en varias repostas.

² Así e todo, nos datos utilizados, a porcentaxe de respostas baleiras (isto é, sen resposta) é de 2,76%.

Sobre os datos das respostas orixinais do cuestionario faise precisa unha intervención, pois cómpre transformalos para que as preguntas poidan ser tratadas cuantitativamente. Este método coñécese en dialectometría como taxación ou tipificación e consiste en representar o esencial da variación dos datos orixinais por medio dun conxunto de unidades taxatorias (*taxats*). A taxación supón unha simplificación dos datos do atlas –que xa sempre son unha representación compendiada da realidade que pretenden analizar–, polo tanto debe realizarse con criterios coherentes que permitan reflectir do mellor xeito posible a variación real observada.

A selección e tipificación das preguntas fíxose seguindo criterios fonéticos, morfolóxicos e léxicos. As 100 preguntas de tipo fonético seleccionáronse tendo en conta as categorías clásicas da fonética histórica e tamén algunhas pautas tradicionais dos estudos de fonética galega. Nas 121 preguntas correspondentes á morfoloxía combináronse morfoloxía nominal e morfoloxía verbal. Aínda que non resulta doado xerbrar neste ámbito fonética e morfoloxía, para a tipificación das respostas atendeuse a variacións que afectan ás formas nominais e ás verbais e que se adoitan considerar de tipo morfolóxico. Nas 100 preguntas de léxico distinguíronse os tipos atendendo aos lemas léxicos etimolóxicos.

PREGUNTAS							
<i>noz</i>		<i>sol</i>		<i>cantade</i>		<i>ti/tu</i>	
RESPOSTAS	TAXATS	RESPOSTAS	TAXATS	RESPOSTAS	TAXATS	RESPOSTAS	TAXATS
noz	1	[ʃɔl]	1	cantade	1	tí	1
nuez		[sɔl]		cantade	2	tu	2
concho	2	[ʃɔl]		cantai	3		
croucho		[s ^θ ɔl]		cantei			
carolo	3	[sol]	2				
carola							

Cadro 2. Exemplos de taxación de catro preguntas do *ALGa*

No cadro 2 móstranse catro exemplos dos resultados da tipificación das respostas. As preguntas corresponden ás tres grandes agrupacións consideradas: léxico (*noz*), fonética (*sol*) e morfoloxía (*cantade* e *tu/ti*). A variación das respostas orixinais tipificouse tendo en conta o trazo que se pretendía analizar. Na pregunta de léxico atendeuse aos étimos das raíces; na pregunta de fonética só se tivo en conta a abertura da vogal tónica; e nas dúas preguntas de morfoloxía consideráronse tra-

zos de carácter morfolóxico. Para cada punto só se valorou a resposta que no cuestionario figura como resposta primeira, xa que o programa estatístico utilizado só permite a introdución dunha única resposta por pregunta para cada punto.

A matriz de datos resultante (321 preguntas x 167 puntos) tratouse con procedementos estatísticos para obter os índices de similaridade e de distancia, empregando en primeira instancia o índice relativo de similaridade e o índice ponderado de similaridade. O índice relativo de similaridade trata a todos os taxats estudados do mesmo xeito, mentres que o índice ponderado de similaridade confire un maior peso aos taxats menos extendidos e axuda a descubrir aqueles trazos lingüísticos merecentes dun tratamento máis atento, ben por seren arcaísmos ben por seren innovacións. Sobre estes índices calculáronse as novas matrices de distancia e de similaridade que nos permiten coñecer, outra vez mediante métodos estatísticos, as relacións de distancia lingüística entre os puntos do territorio analizados. A matriz de similaridade dá conta das semellanzas ou das diverxencias dos diferentes obxectos estudados con relación ás variables analizadas e, trasladada á xeolingüística, serve para coñecer as localidades que máis se semellan polos seus trazos lingüísticos e tamén aquelas que máis se distancian.

Os datos desta matriz poden ser analizados empregando outras ferramentas estatísticas complexas que nos deixan descubrir outro tipo de relacións lingüísticas. O programa empregado, *Visual dialectometry*³, facilita todas estas operacións e cálculos a partir da introdución da información nunha base de datos, e mais permite a representación gráfica e cartográfica dos resultados.

Para a visualización cartográfica dos resultados, a trama de localidades enquisadas converteuse nun mapa de polígonos seguindo os principios da xeometría de Voronoi e Delaunay (Aurenhammer 1991) e do mesmo xeito que se ten feito nas investigacións dialectométricas desde hai xa case trinta anos (Goebel 1981). No mapa de polígonos (mapa coropleto) cada espazo representa a área de influencia de cada

³ Na páxina <http://www.sbg.ac.at/rom/people/proj/dm/vdm/features.html> ofrécese información detallada sobre o programa empregado.

un dos 167 lugares da rede de puntos do *ALGa*⁴. A representación cartográfica facilita a interpretación das análises estatísticas e axuda a comparar os resultados da aplicación do método dialectométrico cos doutros procedementos. Os mapas coropletos permiten converter a variabilidade numérica das matrices de similaridade nunha variabilidade cromática que segue a orde das cores do arco da vella. A representación cartográfica complementábase cos gráficos, histogramas e dendrogramas, que reproducen as frecuencias e as relacións de similaridade entre as localidades estudadas.

Os trece mapas e gráficos que se ofrecen nesta contribución pretenden ser unha mostra da aplicación do método dialectométrico aos datos do *ALGa* e son froito da aplicación do programa *Visual Dialectometry*, deseñado na Universidade de Salzburgo baixo a dirección de Hans Goebel e Edgar Haimerl. As convencións estatísticas e cartográficas adoptadas seguen as propostas realizadas por Hans Goebel nas investigacións dialectométricas sobre o *ALD* (1981) e o *ALF* (2002 e 2003). As análises que se fan a continuación parten da aplicación de tres convencións estatísticas: a sintonía dos valores máximos de distribución de similaridade, a sintonía dos coeficientes de simetría e a clasificación dendrográfica das variedades lingüísticas galegas. A análise dos valores máximos de distribución de similaridade permite identificar as zonas dialectais máis homoxéneas, que acostuman tamén ser aquelas que nalgún momento foron emisoras de influencias.

As figuras 1, 2, 3, 4, 6, 7, 8, 10 e 13 dan conta dos resultados de aplicación desta análise a datos fonéticos, morfolóxicos e léxicos. O coeficiente de simetría que se analiza nas figuras 3, 5, 9 e 12 é un índice estatístico que na súa aplicación dialectométrica axuda a atopar as zonas de forte compromiso lingüístico, máis conservadores, e as zonas de interacción e cambio. Esta análise fundaméntase no cálculo das medias de aritméticas de simetría e na distribución das frecuencias ao redor da media. A figura 13 mostra os resultados da análise

⁴ Nos mapas que se empregan neste traballo aparecen tres puntos engadidos coloreados en gris que non corresponden con ningunha das localidades do *ALGa*. Serven como puntos virtuais con que facer comparación entre as variedades galegas e outras variedades externas ás localidades representadas (variedade estándar, variedades non galegas, etc.).

de conglomerados aplicada aos datos dialectais. Este procedemento axuda a descubrir como se agrupan as variedades xeolingüísticas en clases e tamén a identificar as áreas que se mostran máis distanciadas e singulares.

INTERPRETACIÓN DAS FIGURAS 1, 2 E 3: SIMILARIDADE MÁXIMA E COEFICIENTE DE SIMETRÍA NOS DATOS FONÉTICOS

A figura 1 mostra os resultados da análise do índice de similaridade máxima atendendo só aos datos fonéticos; as puntuacións obtéñense aplicando un índice relativo de identidade, que trata todos os trazos por igual, sen privilexiar os menos difundidos e por iso máis característicos e individualizadores dun territorio.

O mapa mostra que o índice máximo de similaridade logrado por un ou varios obxectos vermellos é do 97,00%, o que implica que ningún lugar de enquisa mostra unha identificación total cos trazos fonéticos doutro. De todos os xeitos, os valores son moi altos: os 33 lugares tinguidos de vermello lograron un índice máximo de similaridade que vai do 94,32 ao 97,00%; estes 33 puntos conforman a quinta parte do total de 167 puntos do *ALGa* (exactamente, o 19,76%), e séguenlles a moi pouca distancia os 74 puntos tinguidos de laranxa (o 44,31% dos puntos do *ALGa*), que acadaron índices de similaridade máxima que van do 91,64 ao 94,32%.

De feito, como mostran o histograma e a curva gausiana de simetría, a media aritmética de similaridade é moi alta (91,65%) e son maioría os puntos que, así a todo, se sitúan á dereita do cumio, con intervalos moi curtos; pola contra, á súa esquerda vese un amplo campo de dispersión respecto da media, cuns poucos lugares moi afastados dela. Os 56 puntos tinguidos de verde (33,53% do total) acadaron índices máximos de similaridade que van do 83,06 ao 91,64%, distribuídos en dous treitos onde son ampla maioría os de similaridade máis alta; en última instancia, o 97,6% dos puntos do *ALGa* (i.e., 163 de 167) acadou índices de similaridade máxima que van do 83,27% ao 97,92%. Finalmente, só catro puntos se sitúan no treito máis baixo, que así a todo vai do 74,49 ao 83,06%.

Estes altos índices de similaridade son aínda máis rechamantes desde que sabemos que as preguntas que se seleccionan para a elaboración dos cuestionarios de dialectoloxía están xeralmente concibidas para detectar ou comprobar diferenzas, de modo que se descartan as partes do léxico ou da gramática que carecen de interese dialectolóxico porque se supón que son trazos comúns a todo un territorio.

Salientamos que os índices de similaridade fonética máxima son moi altos, pero iso non é en si mesmo garantía dunha grande homoxeneidade: cómpre non esquecer que estes datos só nos informan de que cada un dos puntos considerados garda ese nivel de similaridade con aquel con que comparte máis trazos, pero non nos informan directamente sobre o nivel de similaridade cos máis diferentes. Ora ben, se todos os puntos mostran índices moi elevados é presumible que se deba a que as variantes consideradas non provocan unha grande fractura interna, sobre todo se admitimos que por lóxica as coincidencias se dan sobre todo entre puntos colindantes.

Outro elemento que cómpre sinalar é a distribución das zonas vermellas, todas no norte de Galicia. Como sinala Goebl (1993: 288), estas concentracións de máxima similaridade son cernas dialectais, núcleos onde se dá unha situación de homoxeneidade xeolingüística, rodeados de zonas de transición tipoloxicamente non homoxéneas. A zona vermello-laranxa do noroeste da Coruña parte o galego occidental (cf. figura 13), de maneira que os puntos situados ao norte do Salnés mostran unha clara predominancia cara ás cores quentes, indicativa dunha gran similaridade con outros puntos (presumiblemente veciños), en tanto que os situados ao sur mostran índices máis baixos. A outra cerna vermella sitúase en territorio mindoniense, cunha derrama oriental en territorio lucense, e continúa cara ao sur nesa gran mancha compacta de cor laranxa que ocupa o territorio lucuauriense (cf. figura 13).

Se por contra nos detemos nos niveis máis baixos, atopamos tres dos polígonos azuis marcando a fronteira co leonés, e outro no leste do territorio tudense, sobre o Miño, entre o galego occidental e o central. E os puntos verdes, salvante a zona pontevedresa citada antes, non se organizan en grandes ou pequenas áreas compactas, senón aliñados aos dous lados do gran bloque central, como marcando transicións, respecto do occidental e da estrema oriental.

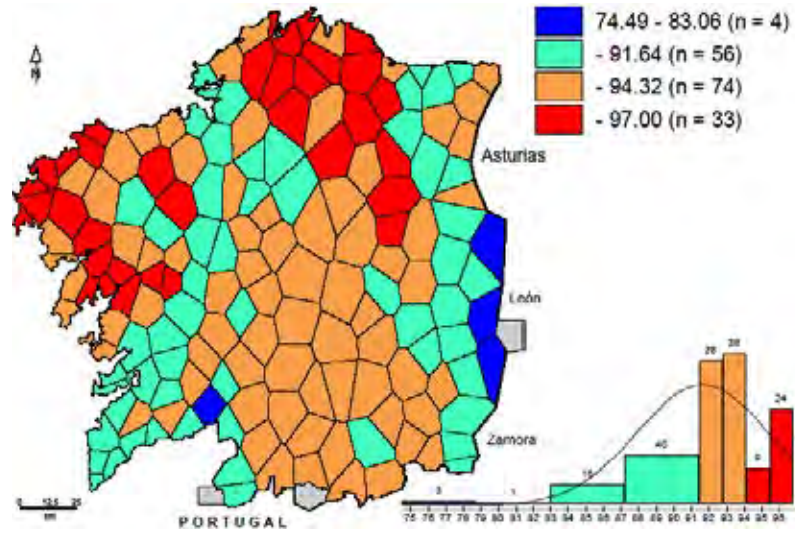


Figura 1. Similaridade máxima para os trazos fonéticos (4 intervalos)

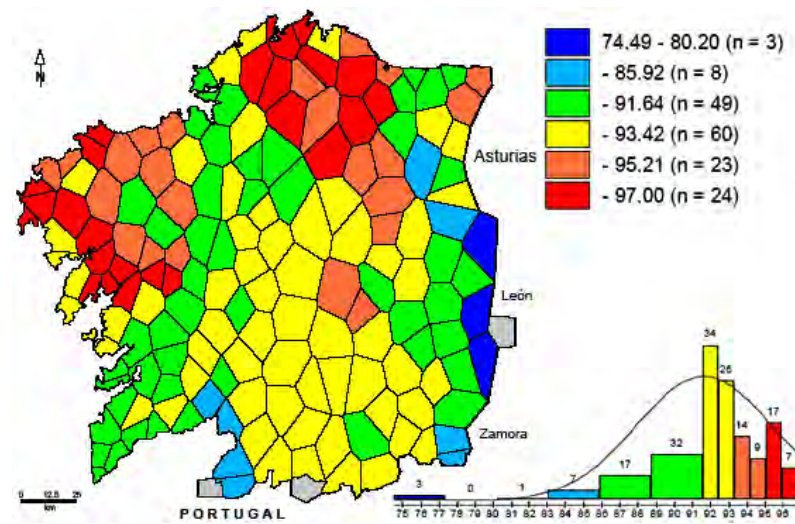


Figura 2. Similaridade máxima para os trazos fonéticos (6 intervalos)

Se cos mesmos datos facemos seis intervalos (figura 2), o debuxo non é moi diferente pero si máis clarificador. Faise máis evidente que as cernas dialectais se sitúan no norte, cunhas áreas de maior intensidade marcadas polos polígonos vermellos, que teñen os seus devalos inmediatos nos polígonos laranxas asociados; a elas podemos engadir esas dúas pequenas áreas laranxas esgazadas, que constitúen á súa vez outros núcleos dialectais con niveis de similaridade menores cós das áreas fisterrá e mindoniense. Faise tamén máis evidente que entre estas dúas áreas de forte presenza no mapa hai unha valgada de transición ben acentuada, na saída natural do interior cara ao golfo ártabro.

Se comparamos as manchas vermellas dos dous mapas (figuras 1 e 2) as diferenzas non son moi grandes, pois permanecen os mesmos polígonos, con moi poucas desercións. Na área occidental rómpese a continuidade na liña costeira que vai de Bergantiños a Fisterra, creando unha valgada na abertura da Terra de Soneira ao mar, e pasa a un chanzo inferior o nexo Cerceda-Oroso, na comarca de Ordes. Na área mindoniense o núcleo permanece tamén no básico, co paso a segundo plano dun polígono interior (Ouro), que se une a outro contíguo marcado xa con laranxa no mapa anterior (Vilalba), e doutro na Mariña lucense (Cervo), que se une así ao de Barreiros; pero o máis salientable é que se distancia esa derrama oriental formada por Meira-Baleira-Baralla, no sistema montañoso que limita a chaira polo leste, que se confirma así como o primeiro chanzo na transición respecto das outras variedades lucenses.

O contraste entre os dous mapas é máis obvio no feito de que o laranxa deixa de ser a cor dominante, pois fóra dos polígonos asociados ás cernas dialectais xa comentadas foi masivamente substituído pola cor amarela. Volvemos salienta o feito de que ao sur dos polígonos ribeiráns do Ulla (Catoira-A Estrada, límite sur da cerna occidental comentada máis arriba) e dese nexo laranxa que permanece en territorio lucuauriense (Bóveda-A Pobra do Brollón), desapareceron as cores máis quentes do mapa e con elas os índices máis altos de similaridade. Se reparamos no que ocorre no norte, vemos que no núcleo occidental se distancian puntos periféricos, sobre todo no seu bordo marítimo SW (Carnota-Muros, Porto do Son-Ribeira, O Grove-Sanxenxo-Cambados, con Moraña); a cor amarela de Vimianzo vai asociada ao cambio a laranxa xa mencionado dos polígonos costeiros veciños, confirmando a

Terra de Soneira como transición entre as falas de Bergantiños e as propiamente fisterrás, e a de Carral forma parte do devalo cara ás Mariñas, antes de comezar o abalo cara ao núcleo mindoniense, agora marcado pola cor amarela dos tres lugares antes tinguidos con laranxa (Monfero no interior, Valdoviño e Cedeira na costa); o carácter de menor homoxeneidade da costa da Mariña lucense, que xa vimos acentuarse en Cervo, compróbase agora coa cor amarela de Viveiro; finalmente, percíbese unha fractura no galego eonaviego, onde os dous puntos nororientais (Coaña e Boal) se manteñen nun chanzo superior aos outros.

Con todo, o máis salientable é ese grande espazo continuo de transición que se crea no centro-sur de Galicia, onde cada polígono comparte un número de trazos moi alto (entre o 91,65% e o 93,42%), mantendo así a todo unha marxe de diferenza sucesiva cos respecto aos puntos colindantes. Un grande espazo que ten no seu interior, en forma de illa, o nexo Bóveda-A Pobra do Brollón (posiblemente por maior homoxeneidade entre ambas) e no extremo NE ese grupo de transición formado por Castro de Rei e os tres polígonos xa mencionados (Meira-Baleira-Baralla); estas son as únicas praias laranxas que contornan ese espazo amarelo, isto é, a única parte do territorio en que os polígonos amarelos supoñen un devalo nesa perda progresiva de homoxeneidade lingüística. No leste, está limitada por unha sucesión de polígonos verdes e azul claro, que marcan a distintos ritmos a distancia lingüística, ata chegar aos tres puntos fronteirizos, onde convencionalmente se sitúa o límite difuso entre galego e leonés. No NW, a sucesión algo descontinua de puntos verdes e amarelos (antes laranxas) marca, en forma de abalo-devalo, a transición entre os dous núcleos mencionados, e en definitiva entre variedades do galego central e do occidental. No S e SW chaman a atención varios aspectos: o polígono illado en azul do mapa da figura 1, aparece agora en azul claro e reforzado polos tres puntos ourensáns da Limia Baixa, que en tantos riscos lingüísticos son mencionados á par dos tudenses. Advírtense índices de similaridade altos na raia seca (exceptuados os mencionados), que contrastan coa cor verde e en parte azul dos ribeiráns na fronteira do Miño, onde a falta de homoxeneidade é patente; o Salnés, con toda a Arousa, forma parte do devalo do grupo coruñés occidental, que chega aos seus niveis máis baixos nas rías de Pontevedra e Vigo, situadas entre ese territorio e a grande mancha central amarela.

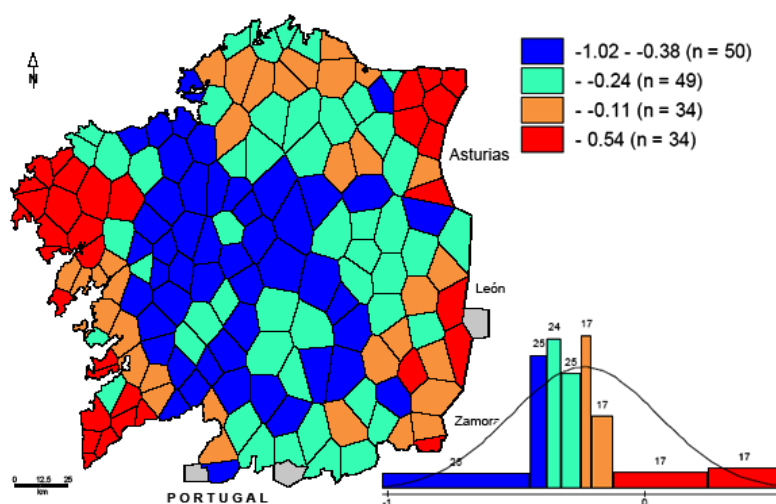


Figura 3. Coeficiente de simetría para os trazos fonéticos (4 intervalos)

A cartografía dos coeficientes de simetría permite visualizar os lugares da rede onde se opera unha maior interacción xeolingüística, de forma que os falantes se atopan en mellores condicións comunicativas por compartiren en maior medida trazos con uns e con outros (polígonos azuis e verdes, nesta orde); en contrapartida, as cores quentes marcan áreas moi conservadoras ou áreas en retroceso, que se pechan en si mesmas e ofrecen maiores dificultades para a interacción con falantes das variedades con trazos máis difundidos. A figura 3 mostra unha imaxe ben recoñecible para os investigadores da estratigrafía lingüística, con ese grande espazo central de interacción lingüística e as zonas máis conservadoras na periferia: unha faixa continua occidental (de forma moi aproximada, a Galicia seseante), que comprende a Limia Baixa, onde salientan o grupo fisterrán no norte e o de Morrazo-Tui no sur; unha banda oriental, co galego eonaviego ben destacado no norte e o grupo que podemos chamar asturicense no sur; o grupo mindoniense, en mellores condicións comunicativas, mantén así a todo ese espiñazo conservador entre a Terra Chá e a primeira liña marítima, con esoutro núcleo SE, nas proximidades da serra de Meira.

INTERPRETACIÓN DAS FIGURAS 4 E 5: SIMILARIDADE MÁXIMA E COEFICIENTE DE SIMETRÍA NOS DATOS MORFOLÓXICOS

Ao analizarmos a distribución de similaridade máxima a través dos datos morfolóxicos, atopámonos con cifras algo máis baixas cás ofrecidas na figura 1, o que indica que se comparten máis trazos fónicos ca morfolóxicos. Esta percepción baséase en que a media aritmética é algo menor (90,31%), son menos os polígonos que se sitúan nos dous chanzos máis altos de similaridade morfolóxica e os lugares de menor similaridade presentan índices moito máis baixos. Ademais os intervalos entre os dous segmentos máis baixos da escala (verde e azul) son de 18,05 valores na morfoloxía, fronte aos só 8,58 do mapa de similaridade fonética, o que apunta tamén a unha maior variación morfolóxica ca fonética.

Hai así a todo un nivel de coincidencia formal moi alto, que se evidencia na alta media mencionada e en que seguen sendo ampla maioría os polígonos situados á dereita na curva en campá; ademais, o 96,41% dos puntos presenta un grao de similaridade superior ao 81%. Pódese ponderar aínda máis esta coincidencia de base ao constataremos que ningún polígono se atopa por baixo do 50%, e os dous de menor similaridade volven ser dous lugares situados sobre o límite oriental (Candín e Pombriego).

O mapa da figura 4, que cartografa a máxima similaridade morfolóxica en catro intervalos, vale para ilustrar esa ampla marxe de coincidencia. Se sumamos os 27 puntos tinguidos con vermello (16,17% do total) aos 65 tinguidos con laranxa, que ofrecen niveis de similaridade moi próximos, comprobamos que máis da metade (55,09% dos puntos) mantén niveis de similaridade superiores ao 90,30%. O histograma revela, ademais, que dos 73 puntos marcados con verde no mapa, 69 teñen valores que están por riba do 81% de similaridade, polo que só restan 6 puntos cuns trazos morfolóxicos ben diferenciados: os dous marcados con azul (54,21-72,25%) e 4 dos marcados en verde (72,26-81%).

Ao observarmos as manchas debuxadas no mapa vemos que se confirma a existencia dun grande espazo noroccidental cunha concentración de puntos de alta similaridade, agora máis grande e compacto, que vai da ría de Foz á de Vigo (grosso modo, as dioceses

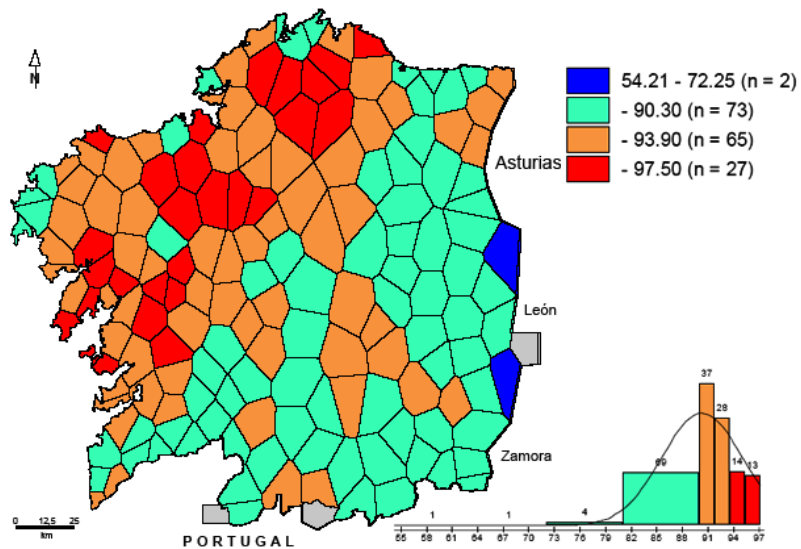


Figura 4. Similaridade máxima para os trazos morfolóxicos (4 intervalos)

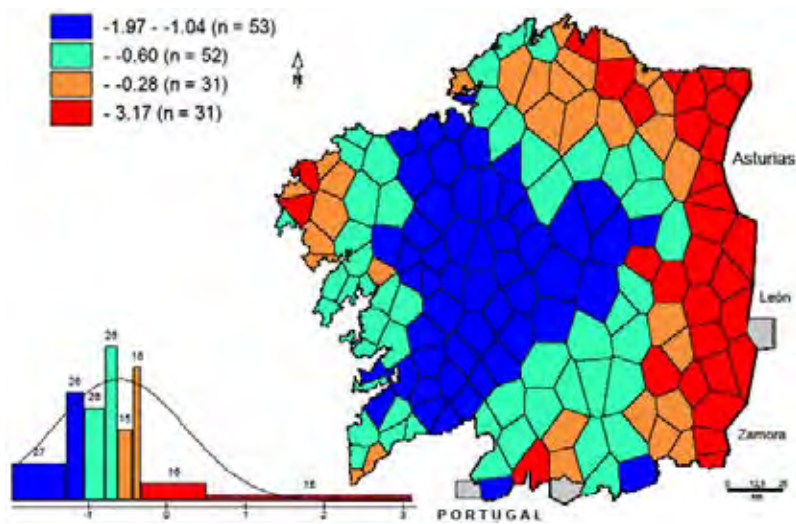


Figura 5. Coeficiente de simetría para os trazos morfolóxicos (4 intervalos)

de Mondoñedo e Santiago), pero unha esculca máis demorada permite observar algunhas diferencias significativas respecto do mapa de fonética correspondente (figura 1). A cerna dialectal mindoniense concentra os polígonos vermellos nun espazo interior centro-occidental, entre as serras da Capelada e o Xistral, de forma que a sucesión de polígonos de menor homoxeneidade que evidencia a transición cara ás variedades orientais, e ao galego eonaviego de forma singular, comeza nun meridiano situado moito máis ao oeste.

No galego occidental obsérvase unha disgregación da concentración de polígonos vermellos e un desprazamento importante na súa localización: en Bergantiños só resta un punto e, sobre todo, na área fisterrá só manteñen os niveis máis elevados os polígonos da parte sur, aos que se engaden outros costeiros máis meridionais situados na Arousa ou nas súas proximidades; xunto a este desprazamento cara ao sur obsérvase outro maior cara ao leste, coa ampliación de áreas vermellas antes representadas de forma máis testemuñal, pois ao nexa Cerceda-Oroso únense agora polo leste e nordés outros cinco polígonos e ao nexa formado polos dous puntos da Estrada engádense un polígono polo norte e dous polo sur. Estas dúas concentracións vermellas, contiguas na disposición norte-sur, dispóñense ao leste da depresión meridiana, no territorio que nos estudos de dialectoloxía se reconece como límite do galego occidental fronte ao central; cómpre engadir que na cerna situada ao norte se ultrapasa esa isoglosa. Nese grande espazo vermello e laranxa singularízanse pola maior distancia morfolóxica pequenas áreas periféricas, entre as que salientan a Costa da Morte e a comarca de Ortegá, que en cambio mostran índices altísimos de similaridade fonética.

Do outro lado desa diagonal imaxinaria Foz-Baiona, predomina o verde, indicando que a homoxeneidade de base é menor na morfoloxía ca na fonética. Faltan absolutamente os polígonos vermellos, en consonancia co xa comentado para a figura 1, pero advírtense algunhas concentracións de polígonos laranxas: unha configuración similar do territorio eonaviego, unha nova ordenación do espazo tudense, parcelas descontínuas da grande mancha auriense observable naquel. Dito doutro xeito, os trazos compartidos no territorio eonaviego, que singularizan esta cerna fronte ás variedades situadas ao seu oeste, separadas por bandas de transición máis ou menos amplas,

pertencen por igual á fonética e á morfoloxía; en territorio tudense, os trazos fonéticos dan índices altos nuns puntos e os morfolóxicos noutros, polo que ambos teñen similar importancia, complementaria, para a comprensión da variación interna nese espazo; en cambio, os polígonos do territorio lucuauriense, tomados un a un, mostran unha maior similaridade fónica ca morfolóxica, de onde se deduce que os trazos máis interesantes desde o punto de vista da variación diatópica pertencen ao ámbito da morfoloxía.

A figura 5, que ilustra en catro intervalos a simetría dos datos morfolóxicos, é tamén moi elocuente. Ese grande espazo central ininterrompido tinguido de azul e bordeado de verde mostra unha gran capacidade de interacción lingüística que atangue a maior parte do territorio, mostrando que os trazos que diferencian uns puntos doutros son amplamente compartidos. Salienta ese grande espazo vermello, tecido de polígonos que mostran variedades conservadoras ou con trazos particulares, pouco compartidos por outros puntos veciños ou distantes; o tránsito cara ás variedades asturianas ou leonesas preséntase, como mostra o mapa, con índices de simetría moi altos respecto do resto dos polígonos do noso territorio. O galego oriental, o galego mindoniense e o galego fisterrán revélanse, neste mapa, como as áreas que manteñen máis características propias, menos compartidas con outros puntos e que por iso interaccionan peor coas restantes variedades, o que semella deixalas nunha posición de maior debilidade fronte á extensión previsible dos trazos asumidos e apoiados por maiores continxentes de falantes. O mesmo se pode dicir das pequenas illas de cores quentes espaxadas na periferia.

INTERPRETACIÓN DAS FIGURAS 6 E 7: SIMILARIDADE MÁXIMA NOS DATOS DE MORFOLOXÍA VERBAL E DE MORFOLOXÍA NOMINAL

Ata agora estivemos considerando conxuntamente todos os datos morfolóxicos. Ora ben, tendo en conta o distinto peso relativo concedido á morfoloxía nominal e á morfoloxía verbal nos estudos dialectolóxicos galegos ata o presente, e máis en particular no establecemento dos tres grandes bloques (occidental / central / oriental) e

nas súas compartimentacións internas, pareceunos conveniente xebrar os datos e cotexar ambos os grupos (figuras 6 e 7). Nunha primeira aproximación observamos as diferenzas entre as porcentaxes de similaridade da morfoloxía nominal e as da verbal, comparándoas á súa vez coas cifras contempladas para o conxunto na exposición anterior. Os datos son ilustrativos por si mesmos:

MORFOLOXÍA NOMINAL			MORFOLOXÍA VERBAL			TOTAL MORFOLOXÍA		
INTERVALOS	% LUGARES		INTERVALOS	% LUGARES		INTERVALOS	% LUGARES	
97,28-100%	29,34%	52,69%	93,81-98,51%	14,37%	55,09%	93,91-97,50%	16,17%	55,09%
94,55-97,27%	23,35%		89,11-93,80%	40,72%		90,31-93,90%	38,92%	
76-94,54%	46,11%	47,31%	74,73-89,10%	43,71%	44,91%	72,26-90,30%	43,71%	44,91%
57,45-75,99%	1,20%		60,34-74,72%	1,20%		54,21-72,25%	1,20%	

Cadro 3. Índices de similaridade na morfoloxía nominal e na morfoloxía verbal

No cadro 3 salienta a amplísima coincidencia na morfoloxía nominal, pois preto da terceira parte dos puntos (29,34%) está por riba do 97,28 % de similaridade, isto é, por riba da maior parte dos puntos de similaridade máxima derivada da morfoloxía verbal (segundo a gráfica, só 4 destes se sitúan na contorna do 97%); engadamos que 8 deses 49 puntos rondan o 100% (cf. histograma da figura 6). Séguelle un nutrido grupo (39 polígonos) que se sitúa no segundo chanzo e que, unido ao anterior, fai que o 52,69% dos puntos enquisados se sitúe nuns índices de similaridade na morfoloxía nominal iguais ou superiores aos 24 puntos de máxima similaridade na morfoloxía verbal.

Visto doutro xeito, un índice de similaridade entre o 93,81% e o 94,54% na morfoloxía nominal está por baixo da media e tinguido de cor verde no mapa, mentres que na morfoloxía verbal eses mesmos índices pertencen ao treito máis alto e por iso lles corresponde a cor vermella. As cifras mostran que os valores máximos dados para o conxunto da morfoloxía son semellantes aos dados para a morfoloxía verbal, con índices de similaridade máis baixos cós da nominal.

Comparando os mapas das figuras 5, 6, e 7 vemos que sempre se destacan eses mesmos dous puntos azuis na extrema oriental co leonés; que a banda oriental ten índices de similaridade máis baixos, de forma máis evidente na morfoloxía verbal ca na nominal, pois esta aparece algo fracturada; e que esa banda oriental de menor similaridade ten unha continuación fragmentada pero perceptible cara a occidente ata atinguir o Atlántico pola área tudense.

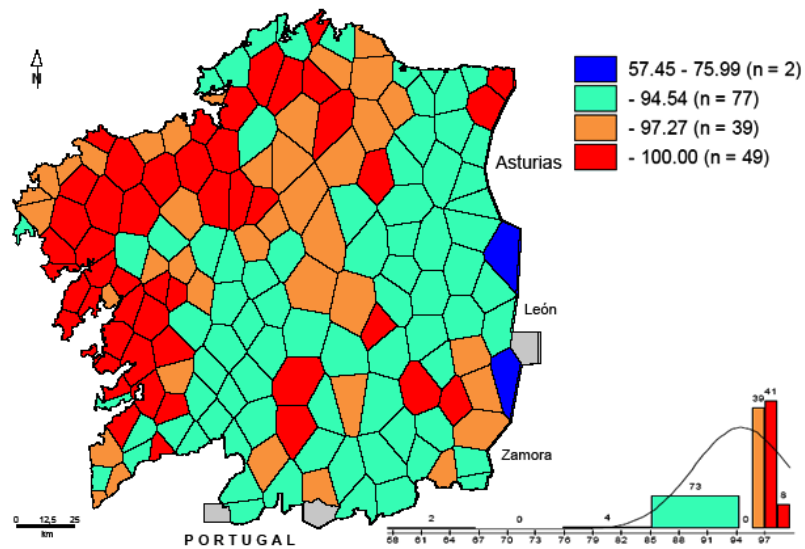


Figura 6. Similaridade máxima para os trazos de morfoloxía nominal (4 intervalos)

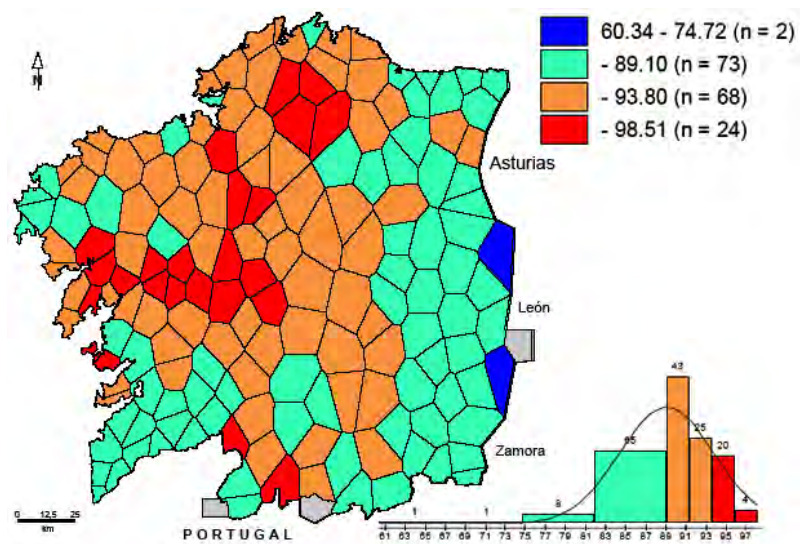


Figura 7. Similaridade máxima para os trazos de morfoloxía verbal (4 intervalos)

Se concentramos a nosa atención nos polígonos situados ao leste da liña Foz-Baiona xa mencionada, onde predomina a cor verde, observamos algunhas diferenzas de interese: considerando separadamente a morfoloxía verbal e, sobre todo, a nominal, observamos que hai polígonos tinguidos de vermello tamén nesta área, que corresponden a lugares onde os niveis máis altos de similaridade nunha parte da morfoloxía se compensan con niveis inferiores á media na outra (polo xeral, moi altos na nominal e baixos na verbal, producindo cor laranxa ou verde no mapa de conxunto); nas variedades do suroeste e do sueste danse niveis moi altos de similaridade na morfoloxía nominal e baixos na verbal, en parte de forma fragmentaria; no espazo máis tipicamente auriense dáse unha situación de altos e baixos, con zonas máis compactas de índices de similaridade relativamente altos na morfoloxía verbal (predominancia de polígonos laranxas contiguos, arredor dunha pequena depresión, que axiña dan paso a unha ampla banda de índice de similaridade inferiores á media ata atinguir o límite oriental), fronte a cumios moi elevados de máxima similaridade na morfoloxía nominal, rodeados de espazos de menor similaridade entre uns e outros (illas vermellas con escasos polígonos laranxa asociados, nun mar de polígonos verdes). Ao oeste da liña Foz-Baiona, por non abundar en consideracións xa feitas ou apuntadas, sinalaremos só o predominio dos polígonos vermellos na morfoloxía nominal e a curiosa concentración de índices de máxima similaridade verbal na conca do Ulla.

INTERPRETACIÓN DAS FIGURAS 8 E 9: SIMILARIDADE MÁXIMA E COEFICIENTE DE SIMETRÍA NOS DATOS MORFOLÓXICOS E FONÉTICOS

Máis arriba sinalamos a similaridade básica entre os datos fonéticos e morfolóxicos, tanto nas porcentaxes alcanzadas polos distintos índices de similaridade máxima, coma na maior cantidade de valores superiores á media e, incluso, pola configuración das manchas que debuxan sobre o territorio os continxentes de polígonos de cada un dos catro intervalos considerados. En todo caso, as diferenzas, que existen, non semellan moito máis significativas cás sinaladas entre a morfoloxía nominal e a verbal; por iso, e porque en última instancia non semellan datos heteroxéneos nin semellan conducir a conclusións diverxentes, pareceunos conveniente considerar todos os estes datos conxuntamente:

FONÉTICA			MORFOLOXÍA			FONÉTICA + MORFOLOXÍA		
INTERVALOS	% LUGARES		INTERVALOS	% LUGARES		INTERVALOS	% LUGARES	
94,33-97,00%	19,76%	64,07%	93,91-97,50%	16,17%	55,09%	93,24-96,36%	11,38%	59,28%
91,65-94,32%	44,31%		90,31-93,90%	38,92%		90,11-93,23%	47,90%	
83,07-91,64%	33,53%	35,93%	72,26-90,30%	43,71%	44,91%	77,01-90,10%	38,92%	40,72%
74,49-83,06%	2,40%		54,21-72,25%	1,20%		63,90-77,00%	1,80%	

Cadro 4. Índices de similaridade na fonética e na morfoloxía

Ao combinar todos estes datos fonéticos e morfolóxicos corrobórase a validez dalgunhas das apreciacións feitas máis arriba (figura 8): son escasos os polígonos situados no treito máis baixo e case todos os que están por baixo da media se sitúan no intervalo máis próximo a esta; doutra banda, son aínda ampla maioría os que superan a media aritmética (agora o 59,28%). Diminúe significativamente o número de polígonos de máxima similaridade, que se concentran en dúas áreas mencionadas de forma recorrente: o grupo mindoniense interior, que mostra índices menores en toda a liña costeira e no seu límite sur, e por baixo da media nas extremas leste e oeste; e o grupo occidental, agora concentrado nun triángulo entre o sur da área fisterrá, a Ulla e o Salnés.

Con todo, na combinación de cores quentes segue percibíndose o mesmo predominio de índices de máxima similaridade na metade noroccidental, onde se dá unha sucesión de polígonos vermellos+laranxas con poucas fracturas verdes, acompañada polo galego eonaviego no recanto nororiental e pola maior parte do territorio lucuauriense; en contraste, as cores frías –e con elas a maior heteroxeneidade nos trazos– predominan na banda oriental e no galego suroccidental, cos puntos de menor similaridade (con todo superiores ao 63,90%) marcando o límite oriental, a transición ao leonés.

A figura 9 ten en conta a distribución do coeficiente de simetría sobre os mesmos datos. A partir da análise estatística ilustra e confirma conclusións ás que é doado chegarmos de maneira impresionística, como observadores atentos da realidade. En primeiro lugar, a existencia desa banda de interacción lingüística entre o bloque central e o occidental, en que os territorios debuxados polas distintas isoglosas se superpoñen, sen dar nunca a sensación de haber espazos marcados por características singulares; está sobre a depresión meridiana e, en definitiva, sobre un grande corredor ancestral norte-sur, que índa hoxe acolle as principais liñas de comunicación do eixe Ferrol-A Coruña-Santiago-Pontevedra-Vigo.

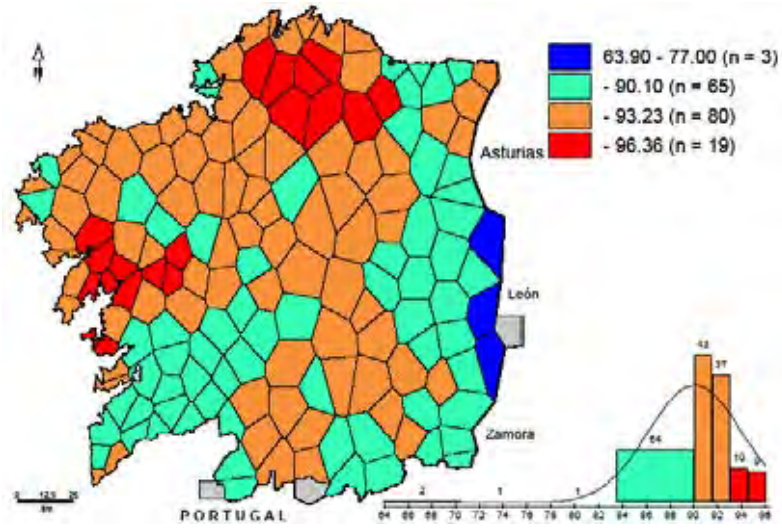


Figura 8. Similaridade máxima para morfoloxía e fonética (4 intervalos)

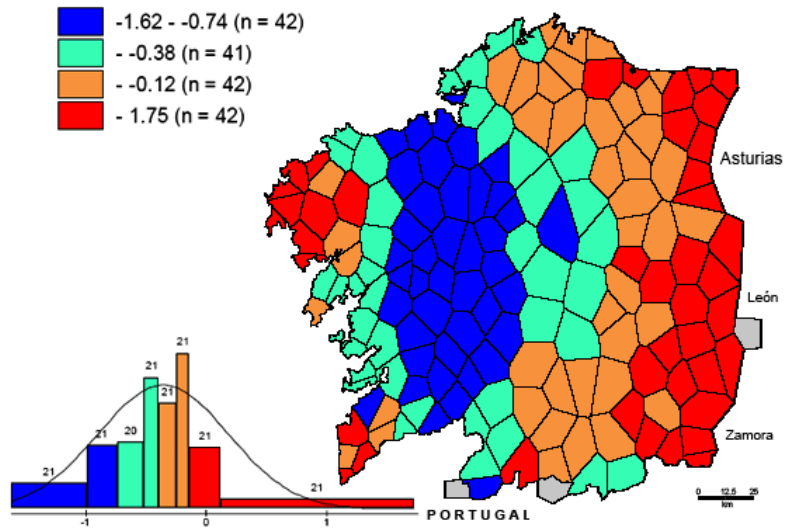


Figura 9. Coeficiente de simetría para morfoloxía e fonética (4 intervalos)

Mirando este mapa non podemos deixar de lembrar a frecuencia con que os procesos de cambio lingüístico, desde hai séculos, progresan deste eixe occidental cara ao leste, máis tamén cara ás áreas periféricas situadas máis ao oeste. En segundo lugar, a existencia desa banda occidental, progresivamente máis marcada conforme avanzamos cara á beiramar, co galego da Costa da Morte e o do Baixo Miño salientados como variedades con trazos de seu máis caracterizadores cós de Bergantiños e das Rías Baixas (Salnés e Morrazo), pero en todo caso estes máis singulares cás variedades lingüísticas do territorio interior situado ás súas costas.

Por último, a percepción de que conforme avanzamos contra o leste imos atopando dentro do galego central trazos cada vez máis propios deste territorio, non amplamente compartidos polo resto, nun continuum W-L, que se van facendo máis numerosos ao nos achegar a esa banda que percorre o límite oriental de norte a sur: é unha área ben coñecida por ser de transición entre o galego central e o oriental, e mais entre o galego e o sistema asturleonés, non compacta senón estruturada en sucesivas variedades dialectais dispostas en dirección N-S, desde o Cantábrico ata Tras-os-Montes (área asturiana, área ancaresa, área oriental central, área zamorana). Os trazos dos puntos situados en toda esa banda oriental non só son de similitude inferior á media, polo tanto menos parecidos aos doutros puntos, senón que son tamén os que mostran características menos compartidas, máis singulares, o que implica tamén que estas son distintas entre uns e outros.

INTERPRETACIÓN DAS FIGURAS 10, 11 E 12: LÉXICO: SIMILARIDADE MÁXIMA E COEFICIENTE DE SIMETRÍA NO DATOS LÉXICOS

Os datos que nos achegan os mapas léxicos dannos índices moi diferentes. A simple comparación do histograma e a curva en campá da figura 10 cos dos anteriores abonda para mostrar que a situación se asemella pouco a todas as descritas ata agora. Como se observa no cadro 5, hai unha especie de inversión nas cifras entre a similaridade máis alta e a máis baixa, e ademais é a primeira vez en que son maioría os polígonos con índices inferiores á media aritmética (50,3%):

FONÉTICA + MORFOLOXÍA			LÉXICO		
INTERVALOS	% LUGARES		INTERVALOS	% LUGARES	
93,24-96,36%	11,38%	59,28%	92,21-99,00%	2,39%	49,70%
90,11-93,23%	47,90%		83,42-91,20%	47,31%	
77,01-90,10%	38,92%	40,72%	79,98-83,41%	37,13%	50,30%
63,90-77,00%	1,80%		76,53-79,97%	13,17%	

Cadro 5. Índices de similaridade en morfoloxía-fonética e en léxico

A observación do mapa da figura 10 mostra tamén que nos atopamos ante unha configuración diferente das manchas que debuxan sobre o territorio os índices de similaridade máxima. Apenas hai polígonos vermellos, só dous nexos que se sitúan así a todo na Galicia noroccidental; por contra, hai unha maior abundancia de polígonos azuis, e non só no límite oriental, pois agora esparéxense nun amplo territorio e ocupan espazos que nos mapas de similaridade fonética e morfolóxica estaban tinguidos coas cores máis quentes. Desapareceu a imaxe reiterada ata agora dunha ampla área vermella e laranxa situada ao norte dunha liña que vai das Rías Baixas á Mariña, e a mancha de cores quentes dominantes sitúase agora ao sur dunha liña que vai de Bergantiños a Valdeorras; ao norte dela, cómpre salientar o grupo mindoniense, sobre as serras que unen a vertente coruñesa e a lucense, así como o lucense central.

Coma no caso da fonética, fixemos unha aproximación aos mesmos datos con 6 intervalos (figura 11), con resultados reveladores por canto se fai máis visible a maior heteroxeneidade léxica xa comentada. Os polígonos de máxima similaridade por riba da media (49,71%) distribúense en tres treitos, mais a práctica totalidade deles está integrada no de índices máis baixos: 2,40% (93,81-99,00%) e 2,40% (88,62-93,80%), fronte a 44,91% (83,42-88,61%). Nos de similaridade máxima máis baixa (50,29%), o número de polígonos é tamén menor segundo se afastan da media, pero non mostran unha diferenza tan acusada entre os tres treitos: 22,75% (81,13-83,41%), 18,56% (78,83-81,12%) e 76,53-78,82% (8,98%). Todo isto percíbese no mapa no predominio dos polígonos amarelos e na práctica eliminación dos laranxas; tamén no dominio dos azuis, ciscados ao leste da depresión meridiana, no canto de moitos verdes no mapa anterior, pero sobre todo interesa salientar o espallamento territorial dos polígonos tinguidos de azul escuro, os de índices de máxima similaridade máis baixos.

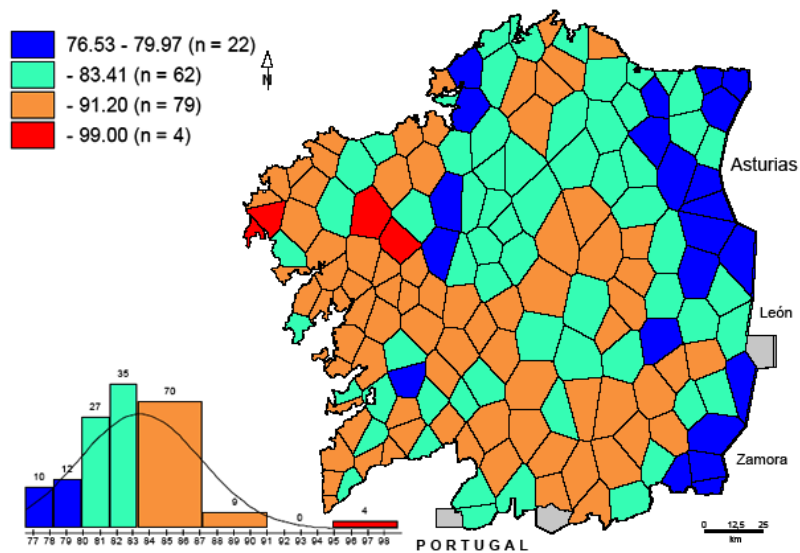


Figura 10. Similaridade máxima para os trazos léxicos (4 intervalos)

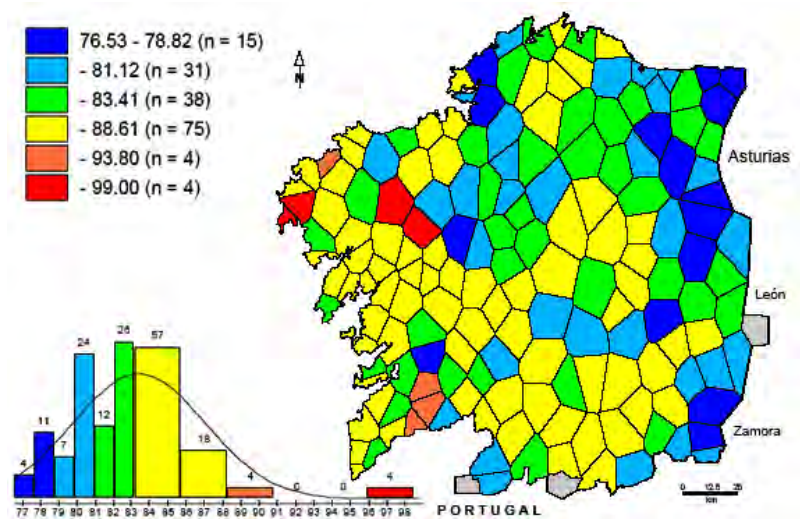


Figura 11. Similaridade máxima para os trazos léxicos (6 intervalos)

A figura 12 mostra a simetría nos datos léxicos. A imaxe non é tan clara na presentación de bandas compactas coma a da figura 9, pero confirma a maior interacción lingüística que se dá na Galicia centro-occidental, creando un espazo de menor peculiaridade léxica local fronte á Galicia centro-oriental. En occidente hai que salientar, con todo, a presenza de pequenos dominios conservadores, de maneira particular, unha vez máis, a Costa da Morte. No oriente destaca a banda vermella oriental, creada a base dunha sucesión de pequenas áreas léxicas para diferentes palabras de pouca difusión, así coma a personalidade dos espazos auriense e, en certo modo, mindoniense.

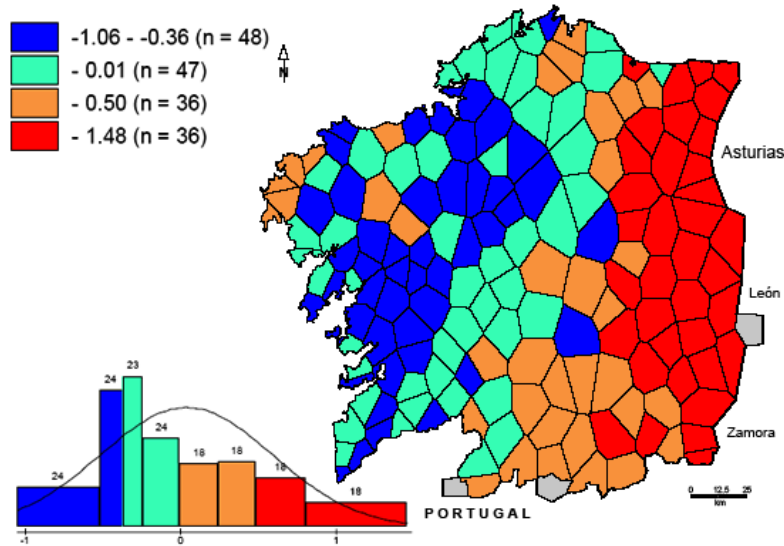


Figura 12. Coeficiente de simetría para os trazos léxicos (4 intervalos)

INTERPRETACIÓN DA FIGURA 13: ANÁLISE DE AGUPAMENTOS PARA TODOS OS TRAZOS LINGÜÍSTICOS

Unha das aplicacións máis interesantes do método cuantitativo é a que permite determinar as agrupacións de variedades lingüísticas que se producen nun territorio lingüístico. Este método denomínase análise de agrupamentos ou conglomerados, unha técnica multivariante que

procura a agrupación de elementos a partir dos trazos que os caracterizan, e que na aplicación dialectométrica vai agrupando as localidades estudadas pola súa proximidade lingüística. A partir da matriz de similaridade comeza o proceso aglomerativo que vai xerando fusionés das localidades que menos se diferencian lingüísticamente; despois de cada fusión xérase unha nova matriz de similaridade que dá lugar a unha nova agrupación. Os algoritmos aglomerativos seguen todos unha mesma estratexia: separan cada elemento nun nodo diferente, calculan a distancia entre os elementos máis próximos e xúntanos nun agrupamento. Os agrupamentos realízanse procurando buscar a maior homoxeneidade intragrupal e a maior distancia intergrupar, polo tanto a distancia entre os elementos dunha mesma clase nunca vai ser superior á distancia entre un elemento desta clase e calquera elemento doutra clase.

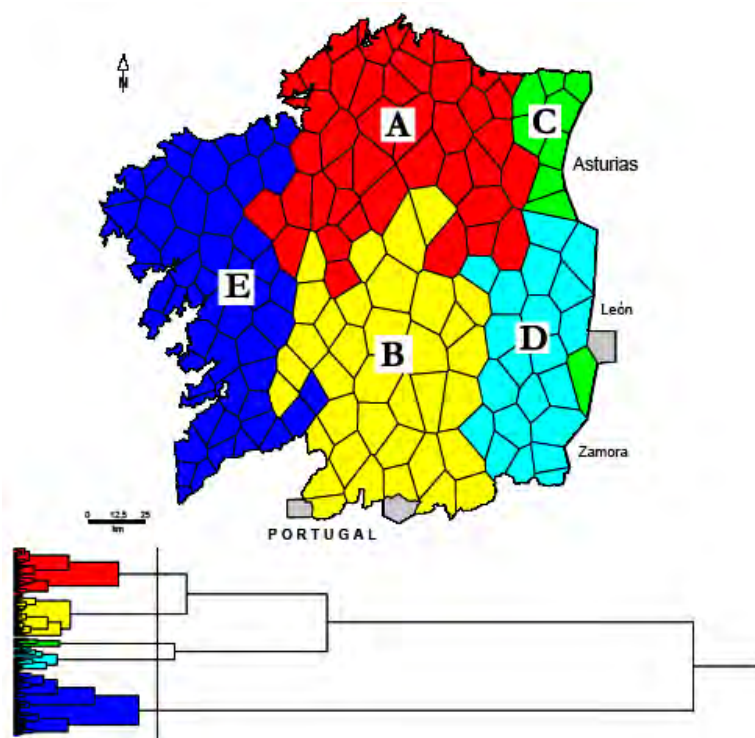


Figura 13. Análise de agrupamentos para o conxunto dos trazos lingüísticos

A representación gráfica desta operación é a árbore xerárquica ou dendrograma. O dendrograma establece unha relación ordenada dos grupos e a lonxitude das súas pólas e gallas é una representación da distancia entre os distintos agrupamentos. O dendrograma aparece acompañado dun mapa coropleto en que se colorearon os polígonos correspondentes aos agrupamentos. As cores do dendrograma e do mapa teñen simplemente unha finalidade cualitativa, e non cuantitativa coma nos mapas precedentes.

O dendrograma da figura 13 estableceuse coa axuda do algoritmo de Ward (Goebel 2000) e mostra as agrupacións que resultan da análise dos 321 trazos lingüísticos para os puntos do *ALGa*. O dendrograma representa en dúas dimensións o proceso de aglomeración dos 167 puntos do *ALGa*: no eixe de ordenadas colócanse as localidades analizadas, que pola cantidade non resultan apreciables individualmente; no eixe de abcisas respreséntanse os niveis sucesivos de agrupamento. A simple vista temos unha idea das agrupacións e da altura a que se produciron.

A lectura debe facerse neste caso de esquerda a dereita. As liñas verticais representan a fusión de agrupamentos: as agrupacións que se encontran máis á esquerda están formadas por localidades que comparten máis trazos comúns; xa que logo, as agrupacións que se producen máis cara á dereita están constituídas por localidades que presentan unha similaridade menor. A representación cartográfica da figura 13 corresponde cun corte do dendrograma que fai visible cinco agrupacións que se marcaron con letras maiúsculas. Partindo da dereita do dendrograma observamos que a árbore se xebra primeiramente na distinción entre unha faixa occidental (E) e o resto do territorio lingüístico (A, B, C, D). O agrupamento da faixa occidental comprende algo máis da metade occidental das provincias de Pontevedra e da Coruña; esta área mostra unha alta homoxeneidade interna e aparece distanciada lingüísticamente do resto do territorio. Se observamos á esquerda da liña de corte vertical, a primeira bifurcación nesta faixa prodúcese no seguinte corte (6 agrupacións) e dando lugar a dúas zonas (norte e sur) pouco distanciadas. Polo tanto, a área occidental, que se corresponde basicamente co territorio que na nosa tradición dialectolóxica se denomina «bloque occidental», aparece como a máis singular e tamén a máis distante das outras variedades xeolingüísticas do galego.

Seguindo a interpretación do dendrograma da figura 13, no camiño cara á esquerda atopamos unha bifurcación que permite distinguir entre unha área central (A e B) e outra oriental (C e D). A agrupación central é a máis extensa e abrangue a maior parte das provincias de Lugo e Ourense e as zonas máis orientais de Pontevedra e da Coruña. A faixa oriental comprende os puntos do leste de Ourense, algúns puntos suroccidentais de Lugo e as localidades do Principado de Asturias e das provincias de León e de Zamora onde se falan variedades galegas. O nivel de homoxeneidade interna destas dúas áreas é moi similar, abonda con reparar en que as seguintes bifurcacións se producen case á mesma altura (A e B, C e D). As agrupacións que resultan do corte no seguinte nivel poñen de relevo as diferenzas entre dúas áreas de semellante extensión na faixa central e outras dúas rexións na faixa oriental.

As coincidencias entre o reparto que resulta da análise de conglomerados e a distribución dialectal realizada por Fernández Rei (1990) resultan evidentes. Como conclusións fundamentais das dúas análises tiramos que as agrupacións dialectais básicas se dan en faixas que van de norte a sur e que distinguen tres áreas: occidental, central e oriental. A análise de conglomerados permítenos ademais apreciar que as variedades da área occidental son as que mostran maiores coincidencias entre si e as que máis se afastan das variedades faladas no resto do territorio. Dentro do bloque oriental obsérvanse tamén dúas agrupacións compactas, unha que reúne os sete puntos de Asturias, un de Lugo e outro de León⁵ (C); e outra constituída polos puntos do terzo oriental de Ourense, surorientais de Lugo e os correspondentes ás localidades das provincias de León e Zamora. No bloque central, un pouco menos compacto có oriental, as variedades congrénganse nun grupo setentrional, algo máis extenso cá área mindoniense e outro grupo meridional que destaca pola súa forte homoxeneidade lingüística.

⁵ O punto de Lugo é L19, Negueira de Muñiz, e o de León Le5, Pombriego. Os trazos da variedade falada en Pombriego coinciden polo xeral máis coas variedades leonesas ca coas galegas.

REMATE

O avance na aplicación de métodos estatísticos no tratamento de grandes cantidades de datos chegou á xeolingüística non hai moitas décadas. A utilización de ferramentas informáticas e de sistemas de información xeográfica axudou a que esta metodoloxía puidese ser aplicada sen moita dificultade na manipulación dos materiais que nos ofrecen os atlas lingüísticos. A dialectometría naceu da conxunción de intereses entre a xeolingüística e a análise de datos, e como unha ferramenta argumentativa que o dialectólogo sempre debe usar como un importante complemento nas súas reflexións e propostas interpretativas. A pesar dos reparos que se lle apoñen a esta metodoloxía, resulta innegable que a dialectometría pon aos ollos dos investigadores un conxunto de relacións entre as variedades lingüísticas que quedarían agochadas se nos limitásemos ao emprego dos métodos tradicionais.

No ámbito galego aínda son poucas as tentativas de aplicación da dialectometría ao estudo da variación xeolingüística (Pöll 1993 e Sousa 2006). Agardamos que esta metodoloxía se consolide e nos vaia ofrecendo novas achegas que axuden a coñecer mellor a historia e o presente da variación xeolingüística en Galicia. Estas novas contribucións permitirannos confrontar de mellor maneira os datos xeolingüísticos con datos de orixe extralingüística e deste xeito explorar sobre fundamentos máis ricos a complexidade do fenómeno da variación lingüística.

ANEXO

Preguntas empregadas para a análise dialectométrica. Na columna da esquerda (Grupo) aparecen os códigos que se empregaron para marcar os distintos trazos (100 fonética, 101 morfoloxía nominal, 102 morfoloxía verbal e 104 léxico); na columna da dereita colócase o nome co que se identificou cada unha das preguntas

GRUPO	PREGUNTA	GRUPO	PREGUNTA
100	aira/eira	100	pe/pè
100	as vacas	100	péxego/pexego
100	auga/agua	100	piollo/pollo
100	baixar/baxar	100	pobre/probe
100	bicho/becho	100	pola/ponla
100	bon/bòn	100	ponte/pònte
100	broa/brona	100	posto/pòsto
100	cabalo/caballo	100	pozo/puzo
100	cabra/craba	100	quebrar/crebrar
100	cabrito/cabirto	100	queixo/queixo (sonor)
100	caixa/caxa	100	quente/calente
100	camiño/camín	100	ra/ran
100	casa/caSa (sonor)	100	rubio/roibo
100	catro/cuatro	100	s- sono
100	ceceo	100	saraiba/sarabia
100	ceo/seo	100	seo/ceo
100	chao(nas)/chan	100	sogro/sògro
100	coiro	100	sol
100	conca/cunca	100	tempo/tèmpo
100	concha/cuncha	100	toupa/teupa/tiopa
100	croio/coio	100	trencha/trincha
100	cruz/crus/cru	100	troita
100	dente/dènte	100	uce/urce
100	dereito/dreito	100	uña/unlla
100	desde/derde/dehde	100	verán/verao
100	-doira	100	verdade/verdá/verdai
100	entroido	100	voz/vòz
100	entroido/entroido	100	xabón/xabrón
100	enxoiro	100	xaneiro/janeiro/saneiro
100	esquecer/esquencer	101	animales/animás
100	faio/faiado	101	ans/ás/ais
100	fame/fòme	101	aquel/aquil
100	filloa/fillò	101	aquela/aquèla
100	flor	101	aqueles/aquelos

100	fol/fole	101	aquelo/aquilo
100	folla/fòlla	101	boa
100	fungueiro/fumeiro	101	bos
100	g/gh	101	cadanseu
100	gando/gado	101	caodrís
100	gardar/guardar	101	cánsaa/cánsai
100	gn/ngn/nk	101	carpíns
100	hora/hòra	101	chea/chía
100	hoxe/hose (despala.)	101	cheo/chen
100	hoxe/huxe	101	colérono/colléronno
100	lan/la	101	comigo/comingo
100	leite/lleite	101	convosco/con vós
100	lingua	101	eixe
100	mallo/manllo	101	el/il
100	man/mao	101	ela/èla
100	mañán/mañá	101	este/iste
100	mazán/mazá	101	estes/estos
100	meda/mèda	101	esto/isto
100	medo/mèdo	101	estrume
100	mel/mèl	101	fume
100	merlo/melro	101	guapísima
100	mesa/mèsa	101	irmá/irmán
100	moa/mola	101	irmán/irmao
100	moito	101	irmáns/irmaos
100	muller/mullèr	101	ladroa
100	muxir/munxir	101	lambona
100	neve/nève	101	lle/ñe/le
100	niño/nío	101	llo/llelo
100	noite	101	-los nenos/os nenos
100	novelo/novèlo	101	maceira/aira/ira
100	novo	101	maceira/ro
100	noz/nos	101	meu
100	ollo/òllo	101	miña/mía
100	padriño/padrín	101	o río/el ríu
100	palatalización a átono	101	ó/ao/al
100	palatalización a tónico	101	o/lo (pron)
100	pantalón/òn	101	pantalóns/-ós/-ois
101	te-che/che/te	102	trouxen/truien
101	teu/tou	102	tuyen/tiven/teven
101	tu/ti	102	valer/valir
101	ua(nas)/unha	102	vallo/valio/valgo
101	uns/unhos	102	veu/veo/viu

101	veciña/vecía	102	viñeches/viche
101	veciño/vecín	103	alcume
101	veciños/vecíus	103	andar a gatas
101	vímolo/slo	103	angazo
101	viuna/viua	103	año
101	vós/vosoutros	103	bagazo
102	agradeces/agradèces	103	berce
102	caer/caír	103	bolboreta
102	cantabamos/cantábamos	103	borra
102	cantábamos/cantábanos	103	broa
102	cantade/cantande/cantai	103	cabrito
102	cantades/cantandes/cantais	103	cadeado
102	cantanon/cantaron	103	cadela
102	cantei/cantín	103	canteiro
102	cantemos/cantèmos	103	carreteiro
102	-che/-ches	103	casarío
102	collamos/cóllamos	103	cazapo
102	colle/culle	103	ceo da boca
102	collemos/collèmos	103	chepa
102	collemos/collèmos/collimos	103	choca
102	colleu/colliu	103	cincha
102	collí/collín	103	códea de pan
102	colliche/colleche	103	corte
102	dea/día/de/deña	103	cribo
102	deches/diches	103	embigo
102	deu/dou/diu	103	érbedo
102	dicir/decir/decir	103	espiga
102	din/dei/dein/den	103	espiñazo
102	dis/dices	103	espirrar
102	dou/don	103	estadullo
102	está/estén	103	faiado
102	estar/tar	103	ferrador
102	este/esteña/estxa/estía	103	fervenza
102	facer/faguer/faer/fer	103	fociño
102	fago/faigo/faio	103	fol
102	fas/fais/faes	103	frouma
102	fixen/figuen/fecen	103	garavanzo
102	fumos/fomos	103	hombró
102	habemos/hemos	103	hórreo
102	haxa/haixa/haba/haiga	103	labio
102	houben/huben	103	legaña
102	iba/ía/iña	103	luntura

102	ides/vades	103	levar ó cabaliño
102	oír/ouvir/ouguir	103	lintel
102	partiu/parteu	103	lóstrego
102	podoo/poido/podio	103	lousa
102	pomos/poñemos	103	lousas
102	porei	103	lúa nova
102	pos/pois	103	machada
102	puiden/puden	103	mallo
102	puña	103	mandíbula
102	puxen/puñen	103	marco
102	queira/quira/	103	mazaroca
102	queres/ques	103	médula
102	sallo/salo/saio	103	meniña do ollo
102	sea/seña/seXa	103	moega
102	sei/sein	103	moucho
102	son/sòn	103	munguir
102	souben/soupen	103	neto
102	-stes/-stedes/-chedes	103	noz
102	tendes/tedes	103	orballo
102	teñen/ten/tein	103	orneo
102	ter/tir	103	osa maior
102	tès/tes/teis/tèis	103	pardao
102	tiña/tía	103	pé do muíño
102	traer/traguer/trer/traír	103	pega
102	traigo/traio	103	pingo
103	pita cega	103	tixola
103	pitorrei	103	tobo
103	porco	103	truco
103	porta do forno	103	vaca saída
103	poxa	103	vacada
103	queixo	103	valo do carreiro
103	rebaño	103	vencello
103	rego	103	ver
103	rella	103	vertedeira
103	remuíño	103	vía láctea
103	roxón	103	vinculeiro
103	salamántiga	103	viruela
103	samesuga	103	xenro
103	saramago	103	xudía
103	soro	103	xugo de bois
103	terreo pouco soleado	103	xunguir
103	timón	103	zoca